数理解析研究所講究録
680巻 1989年 285-294

上下限制約をもつ非線形最小二乗問題に対する Affine scaling を用いた ε－近似解法

愛知大　経営　相良信子（Nobuko SAGARA）
京都大　工学　福島雅夫（MASAO Fukushima）

非線形最小二乗問題の解を求める論文は沢山あるが 制約条件をもつ問題を取り扱った論文はあまり多くない。ここでは 次の問題の数値解法を提案する。

$$(P_0)\qquad \min \frac{1}{2}\|F(x)\|^2$$

$$\text{s.t.}\quad l \leqq x \leqq u \qquad x \in R^n$$

ただし $f(x) = \frac{1}{2}\|F(x)\|^2 = \frac{1}{2}\sum_{i=1}^{m} F_i^2(x)$

$F_i \in C^2$　　l と u は x の下限と上限の n 次元ベクトル

次に展開する (P_0) の新解法の特徴は：

① 反復法　② 上下限制約の特徴を考えた trust region法。（以下 t.r. 法と書く）通常の t.r. 法とは異なる。③ 内点法

◦ 本稿で提案する方法の基本的な考え方

通常の t.r. 法 [2, 3, 4, 5] においては 現在点を含む適当な領域（この領域のことを t.r. という）の中で目的関数 $f(x)$ を近似した2次関数を最小化することによって得られる点を

次の反復点とする。関数の性質や領域の大きさにより元の関数に対する近似の精度が異なるので、t. r. 法の収束を保証するように領域の大きさを調整しなければならない。本稿で提案する方法では、常に上下限制約の内部に含まれる点のみを考える。したがって、現在の点 x を含む t. r. として、上下限制約に含まれるような楕円を考え、その中に次の点を定める。よって、現在点 x から次の点 x^+ を決定する修正量を p とすると ……（すなわち 次の点は $x^+ = x + p$ であるが）……（P_0）より次の部分問題を導くことが出来る。

$$(P_1) \qquad \min_{p \in R^n} \frac{1}{2} \| F(x) + J(x) p \|^2$$

$$\text{s.t.} \quad \sum_{i=1}^{n} \left(\frac{p_i}{d_i} \right)^2 \leq \Delta^2 \qquad (0 < \Delta < 1)$$

ただし　$d_i = \min\left[(u_i - x_i), (x_i - l_i) \right]$

$J = J(x) = \nabla F(x)$　　$F = F(x) = (F_1(x), \cdots\cdots, F_m(x))^T$

（P_1）を書きかえると次の問題（P_2）が得られる。

$$(P_2) \qquad \min g^T p + \frac{1}{2} p^T B p \qquad \text{s.t.} \ \| Dp \| \leq \Delta$$

ただし　$g = \nabla f(x) = J^T F$　　　$B = J^T J$

$$D = \begin{bmatrix} 1/d_1 & & 0 \\ & \ddots & \\ 0 & & 1/d_n \end{bmatrix}$$

制約条件 $\|Dp\| \leq \Delta$ は現在の点から上下限制約の境界への距離にもとずいた metric のもとで球になっている。この事からこの方法は Affine scaling 法と呼ばれており、LP 問題に

対する新しいアルゴリズムにおいてよく用いられる考え方である。[1]

次に(P_2)で計算したpを用いてよいかは、$x^+=x+p$の目的関数値がxの関数値より充分に減少しているかを判定することが必要。もし、この条件が充されれば$x^+=x+p$として次の反復点x^+が決まる。そうでなければ、現在点xを動かさずにt. r.の半径を小さくして部分問題(P_2)を再び解く。この手続きを繰返すと、必ず目的関数値を減少させるような点が得られる。この様な操作を各反復点で実行することにより、目的関数を単調減少させる点列が出来、その点列が最適解に近づく。

◦ Active set 法に基づいた(P_2)の修正

(P_0)の最適解は上下限制約の境界上にあることが一般である。上に述べたt. r.法を実行すると現在点が境界に近づくにつれて楕円の長軸/短軸$\to\infty$となり楕円の厚みがなくなる。従って、部分問題(P_2)はill-conditionedになりその解を計算する際に数値的不安定性が生ずる。これを避けるために、制約付非線形最適化手法の1つであるActive set法の考え方をとり入れた修正を行なう。まず、境界の内側に$\varepsilon>0$の巾をもたしておいて変数がその領域に入ったとき、その変数はその値で固定して定数とみなす。その結果定まる部分空

に限定して(P_2)を考えれば ill-conditioned ではなくなる。この様にして部分問題(P_2)は次の様に修正される。

(P_3)
$$\min_{p \in R^n} g^T p + \frac{1}{2} p^T B p$$
$$\text{s. t.} \quad \| Dp \| \leqq \Delta, \qquad 0 < \Delta < 1$$

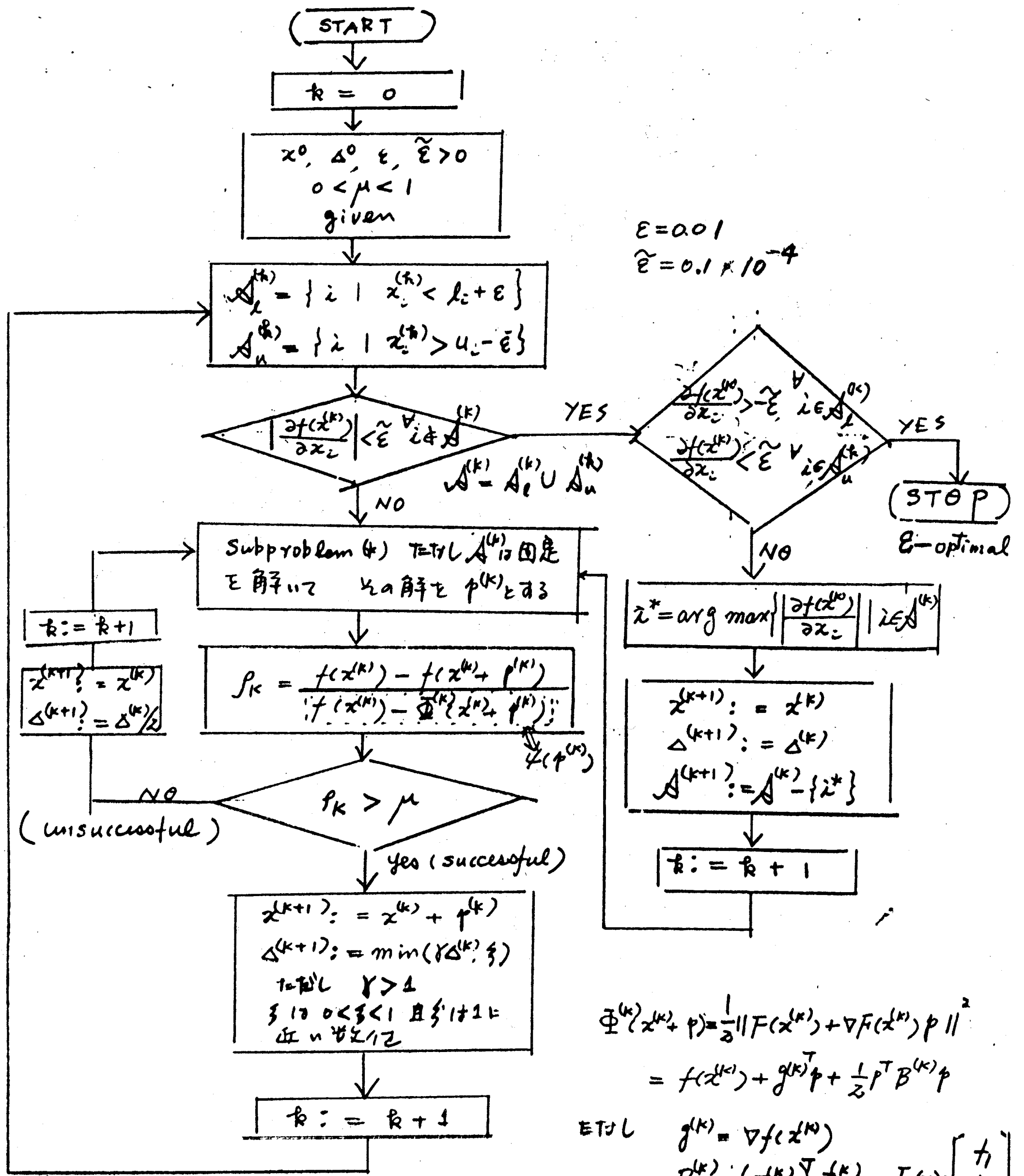
ただし
$$p_i = 0 \quad \text{for } {}^\forall i \in A$$
$$A = A_l \cup A_u$$
$$A_u = \{ i \mid x_i > u_i - \varepsilon \}, \qquad A_l = \{ i \mid x_i < l_i + \varepsilon \}$$

(P_3)は固定されていない変数だけで取り扱うので次元が落ちる。その部分空間においては問題は ill-conditioned にはならず，数値的不安定性は生じない。

(P_3)の解が(P_0)に対する最適性の条件を充せば停止する。そうでなければ，固定された変数を変化させることによって更に目的関数値を減少させることができる。この様な手続きによって(P_0)の最適解の ε－近似解が得られる。この計算の手順を示したのが次頁のフローチャートである。

最後に得られた解の精度は ε に依存する。したがって元の問題の最適解を得るためには修正が必要である。それには，固定されている変数の値をその境界値に直す。すなわち，$l_i < x_i < l_i + \varepsilon$ ならば $x_i = l_i$ にする。そして，それらを固定する。この値を初期値として 固定してしまった変数を除いた部分空間で L. T. 法で解くと最適解を得ることができ

3。



$$\Phi^{(k)}(x^{(k)} + p) = \frac{1}{2} \| F(x^{(k)}) + \nabla F(x^{(k)}) p \|^2$$
$$= f(x^{(k)}) + g^{(k)T} p + \frac{1}{2} p^T B^{(k)} p$$

ただし $g^{(k)} = \nabla f(x^{(k)})$

$B^{(k)} = (J^{(k)})^T J^{(k)}$ $\quad F(x) = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix}$

$J^{(k)} = \nabla F(x^{(k)})$

Subproblem (k)

$$\psi(p) = g^T p + \frac{1}{2} p^T B^{(k)} p \longrightarrow \min_{\| P_k p \| \leq \Delta^{(k)}} \quad (0 < \Delta^{(k)} < 1)$$

where $p_i = 0$ for $\forall i \in \mathcal{A}^{(k)}$ $\quad p \in R^n$

○ <u>新しい手法の計算の動きを見るために</u>

Rosenbrock function を修正した関数[7]

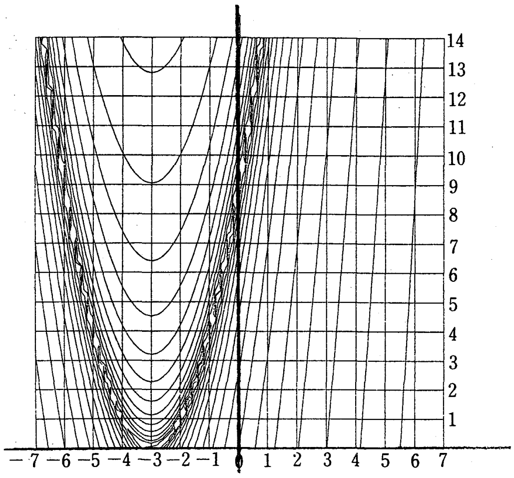
$$F_1(x) = 10.(x_2 - (x_1+3)^2)$$

$$F_2(x) = 2. + x_1$$

について 次の問題を考える。

$$\min_x f(x) = \frac{1}{2}(F_1^2(x) + F_2^2(x)) \quad s.t.\ x \geqq 0,\ x \in R^2$$

この真の解は $x^* = (0,\ 9)^T$

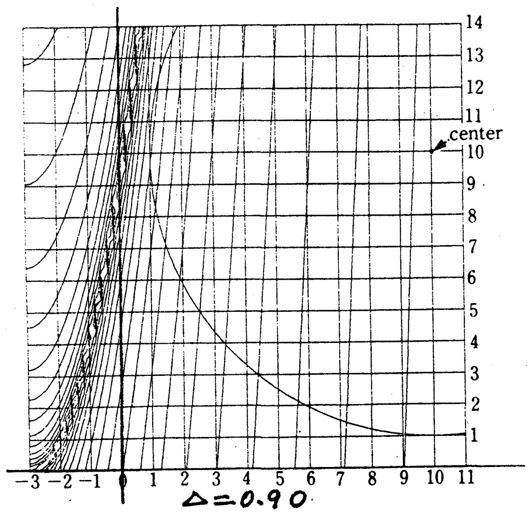


修正した Rosenbrock function

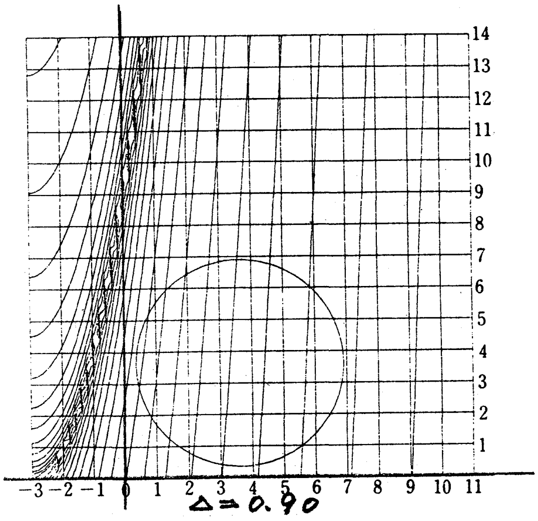
$x^0 = (10.\ 10.)^T$ を初期値としたときの新しい手法の反復点の動きは下記の通り： $\varepsilon = 0.01$

Iter.	x_1	x_2	$\|F\|$
0	10.	10.	1590.
1	3.63	3.63	404.
2	0.51	2.63	97.
3	0.14	4.03	58.
4	0.11	7.57	21.
5	0.0136	9.072	2.0
6	0.1-02	9.0(2)	2.0(2)

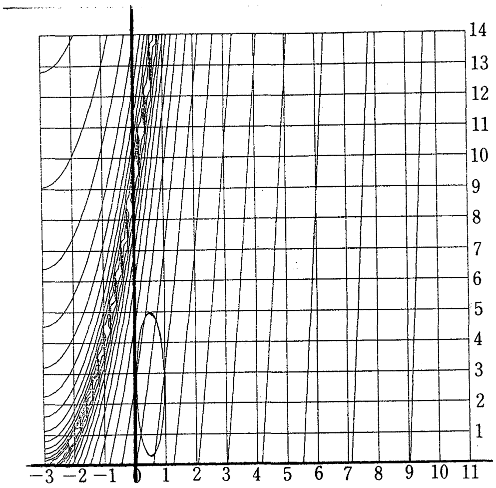
これらの最初のいくつかの反復点における x, r の円を描いたのが次の図である。



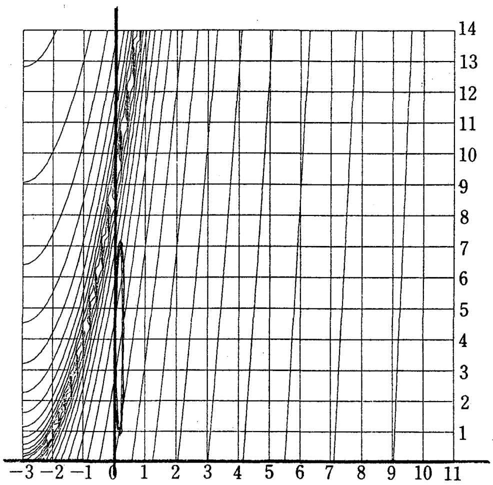
$\Delta = 0.90$



$\Delta = 0.90$



$\Delta = 0.9\dot{9}$



$\Delta = 0.9\dot{9}$

○新しいアルゴリズムの収束について

このアルゴリズムは有限回の反復でε－最適解に到達することが証明できる。

Lemma 1

$$-\psi(p) \geq \frac{1}{2} \min_i |\hat{g}_i| \left\{ \Delta,\ |\hat{g}_i| / \|\hat{J}_i\|^2 \right\}$$

ただし　$\hat{g}_i = x_i \cdot \frac{\partial f}{\partial x_i}$,　$\hat{J}_i = x_i \cdot \frac{\partial F}{\partial x_i}$

このiは $\left| \frac{\partial f(x)}{\partial x_i} \right| \geq \tilde{\varepsilon}$　且　$i \in A$ を充す特定なもの

Lemma 2　　$\Delta \to 0$　は起らない

証明　$\Delta \to 0$ を仮定して $\rho \to 1$ を証明し、Δ の更新の規則に反する矛盾を導く。この証明の途中で lemma 1 を用いる。

Def.　$\rho_K > \mu$（フローチャート参照）を充すとき successful といい、その他のとき、unsuccessful という。

Lemma 3　各反復点で unsuccessful は有限回しか起らない。

証明　Lemma 2 より自明.

Lemma 4

$$f(x) - f(x+p) \geqq \frac{1}{2}\mu \min_i |\hat{g}_i| \{\Delta, |\hat{g}_i| / \|\hat{J}_i\|^2\}$$

ただし　$p_i = 0$　　for $^\forall i \in \mathcal{A}$

Theorem　有限回の反復で ε − 最適解に到達する.

証明　Lemma 2, 3, 4 より導びかれる.

・ 数値実験

新しい手法と従来からの手法である Active set 法[5]（Gauss − Newton）の2つの方法を用いて数値実験を行なった。その結果、変数の個数が大きくなるに従って新しい手法が有効となることが分った。これらの計算は京都大学大型計算センターで行った。

Wrong Wood function[6] $\quad m = 3(n-2)$

$$f_{6i-5}(x) = 10(x_{2i} - x_{2i-1}^2)$$

$$f_{6i-4}(x) = 1 - x_{2i-1}$$

$$f_{6i-3}(x) = 3\sqrt{10}(x_{2i+2} - x_{2i+1}^2)$$

$$f_{6i-2}(x) = 1 - x_{2i+1}$$

$$f_{6i-1}(x) = \sqrt{10}(x_{2i} + x_{2i+2} - 2)$$

$$f_{6i}(x) = \sqrt{10}(x_{2i} - x_{2i+2})$$

$$\min_x f(x) = \frac{1}{2}\sum_{i=1}^{m} f_i(x) \quad \text{s. to} \quad x \geqq l$$

下限lは適当なランダムな正の数値を選び，初期点は$l+\varepsilon$より充分に大きいような点をとった。また，計算で用いたεと$\hat{\varepsilon}$の値は $\varepsilon = 0.001$, $\hat{\varepsilon} = 0.1 \cdot 10^{-3}$ である。この数値実験の結果のCPU time と反復回数は次の表の通り：

n	m	New Method	Active set method
20	54	0.479 (sec)	0.462 (sec)
50	144	14.674	17.757
70	204	46.727	69.731
80	234	78.598	119.137
90	264	118.143	212.276
100	294	161.191	324.306
110	324	223.162	476.646

n	New Method	Active set Method
20	23	30
50	50	72
70	59	102
80	67	117
90	71	132
100	71	147
110	74	162

n	$\dfrac{\text{CPU time for New Method}}{\text{CPU time for Active Set Method}}$	$\dfrac{\text{\# of iter. for New Method}}{\text{\# of iter. for Active Set Method}}$
20	1.036	.766
50	.826	.694
70	.670	.578
80	.659	.572
90	.556	.537
100	.497	.482
110	.468	.456

ε－最適解を補正して最適解を得るために要する計算時間は，nが50以上のときはε－最適解の計算時間の6%から7%位であることが数値実験より分った。従ってnが大きいときは新しい手法が優れていることが分る。

この研究の一部は愛知大学研究助成 C-1 によってなされた．

参考文献

[1] E. R. Barnes: A variation on Karmarkar's algorithm for solving linear programming problems, Math Prog. 36 (1986) 174-182 [2] J.J. Moré and D.C. Sorensen: Newton's Method, ANL-82-8 Feb (1982) [3] J.J. Moré and D. C. Sorensen: Computing a trust region step, SIAM J Sci Stat. Comput. Vol.4 553-572 (1983) [4] J.J. Moré: The Levenberg-Marquardt Algorithm: Implementation and theory, Proceeding of the Dundee Conference on Numerical Analysis, G. A. Watson ed., Springer-Verlag (1978) 105-116 [5] R. Fletcher: Practical Methods of Optimization Vol 1. Vol 2 John Wiley & Sons (1981) [6] Ph. Toint: On large scale nonlinear least squares calculations, SIAM J. Sci. Stat. Comput. Vol. 8 No.3 (1987) [7] 相良: Trust Region法を用いた非負制約をもつ非線形最小二乗問題 愛知大学法経論集第116 昭63年3月